Waters

ANALYSIS OF COMPLEX PROTEIN MIXTURES USING THE ALLIANCE® BIOSEPARATIONS MODULE AND 2D LC/MS/MS

Chris Hughes[1], Therese McKenna[1], Scott Berger[2] and James Langridge[1]
[1]Waters Corporation, MS Technologies Centre, Manchester UK
[2]Waters Corporation, Life Science R&D, Milford, MA, USA

OVERVIEW

In this application note we compare three different analytical strategies for the separation and analysis of a complex *E.coli* cytosolic protein fraction.

The first method involves digestion of the protein mixture and peptide level separation by on-line 2D LC/MS/MS.

The second method involves digestion of the protein complex mixture and off-line peptide separation using the Waters® Alliance® Bioseparations Module, prior to analysis by on-line LC/MS/MS.

The final method used is separation at the protein level, using the Alliance Bioseparations System, followed by trypsin digestion and subsequent 1D LC/MS/MS analysis.

INTRODUCTION

Advances in both HPLC and mass spectrometry instrumentation have allowed the analysis of protein complexes that have not been separated by the established technology of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE).

More specifically, the use of 2D liquid chromatography has gained prominence over the last few years, with the technique being widely adopted by numerous groups.

The most commonly used approach to 2D chromatography usually involves separation of peptide mixtures from a complex biological sample, such as a whole cell lysate or a sub-cellular fraction, which has been digested without any comprehensive protein separation. This 2D peptide separation is usually based on peptide charge, using Strong Cation Exchange (SCX) chromatography followed by separation based on peptide hydrophobicity, using Reversed Phase (RP) chromatography.

Implementation of 2D chromatography for the separation of complex peptide mixtures can be accomplished by either an online or offline method.

The Waters® Alliance® Bioseparations System with the Waters CapLC® and Micromass® Q-Tof Ultima™ API Mass Spectrometer.

In the online experiment, SCX and reversed phase columns are configured on one HPLC system and separation is achieved in a sequential manner where peptides are eluted from the SCX column onto the C_{18} material by increasing salt concentration. This can be achieved using either a linear gradient or by discrete steps of increasing salt concentration. A reversed phase gradient is then performed to elute peptides from the C_{18} material into the mass spectrometer. In the offline 2D experiment, an SCX gradient is run using one HPLC system and peptide fractions are collected for analysis by reversed phase chromatography using a separate HPLC system.

Separation at the protein level can be achieved with the use of the Waters Alliance Bioseparations Module. Fractions from such an HPLC run can be collected and subsequent tryptic digestion of these fractions will result in less complex peptide mixtures that can then be analyzed by nanoscale LC/MS/MS.

In this study a K12 strain derived *E. coli* cytosolic sample has been analyzed using these different strategies and the results compared. The results in the first two strategies are obtained from an analysis where the protein sample was initially digested and then subjected to both on-line and off-line 2D LC/MS/MS.

The final strategy was to separate at the protein level into fractions. These fractions were collected, digested and each subsequently subjected to 1D LC/MS/MS.

METHODS
Sample preparation – Tryptic digestion of *E. coli* sample for on-line and off-line 2D LC/MS/MS.

An aliquot of 160 µL of concentrated *E. coli* cytosol in 20 mM Ammonium Bicarbonate was diluted in 0.1% RapiGest™ SF and 1 mM $CaCl_2$ and digested by adding 160 µg of sequencing grade trypsin (Promega, Madison, WI) in re-suspension buffer. The mixture was digested at 37 °C for 4 hours and then centrifuged at 12,000g for 10 minutes. The supernatant was collected and stored prior to analysis.

2D LC/MS/MS Online Separation (Figure 1)
• HPLC: Waters CapLC® system with Stream Select Module
• Columns: SCX trapping cartridge, 0.020" ID x 20 mm, Symmetry® C_{18} trapping cartridge and Atlantis™ NanoEase™ 75 µM x 100 mm analytical column
• 4 µg sample was injected at a flow rate of 15 µL/min for 6 minutes
• RP Gradient: 5 to 30% acetonitrile in 90 minutes
• Fifteen subsequent salt injections (volume 20 µL), each followed by RP gradient
• Salt concentrations: 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200 and 500 mM KCl
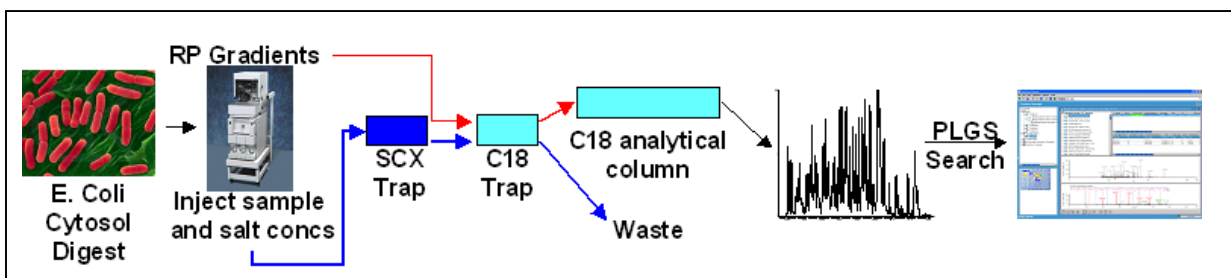All salt solutions made up in 5 mM K_2HPO_4, pH3 in water/acetonitrile (95/5, v/v)

Figure 1. Schematic Representation of On-line 2D LC/MS/MS.

2D LC/MS/MS Offline Separation (Figure 2)

• HPLC: Waters Alliance Bioseparations System

• Column: SCX, 2.1 mm ID x 35 mm long

• Solvent A: 5mM phosphate buffer in 20% acetonitrile

• Solvent B: 1M KCl in solvent A

• Approximately 230 µg of peptide mixture was loaded onto the column

• Gradient: 0 to 27% B over 30 minutes at a flow rate of 100 µL/min

• The eluent was connected to a UV detector and then to a fraction collector into which 300 µL fractions were deposited

• Fractions were diluted four-fold to reduce acetonitrile composition and 10 µL of each fraction was injected onto the nanoscale LC system described later

• RP gradient: 5 to 30% acetonitrile in 90 minutes

Protein prefractionation of the *E. coli* Cytosol using the Waters Alliance Bioseparations System (Figure 3)

• HPLC: Waters Alliance Bioseparations System with 2796 Bioseparations Module and 1525µ Pump both operating at 400 µL/min

• Valve configuration: One SCX first dimension column with two RP second dimension columns

While one of the second dimension columns is being loaded with eluent from the SCX column, the other elutes into the UV or MS detector.

• First dimension column: Waters BioSuite™ SCX column, 4.6 mm ID x 35 mm long

• The AutoBlend™ feature allowed an acetate buffer system to maintain a pH of 5 for the first dimension separation

• Ionic strength adjusted by altering the composition of solvent C (1M NaCl) with respect to solvent D (water)

• Second dimension columns: Waters Symmetry C_4, 2.1 mm ID x 10 mm long

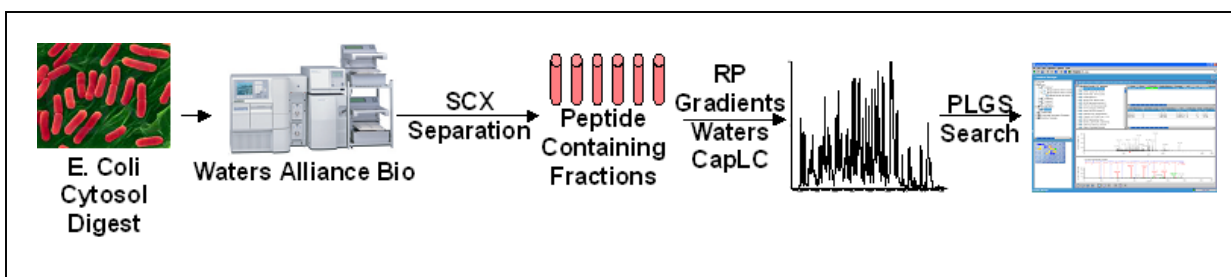• Second dimension solvents: 0.1% aqueous formic acid and 0.1% formic acid in acetonitrile

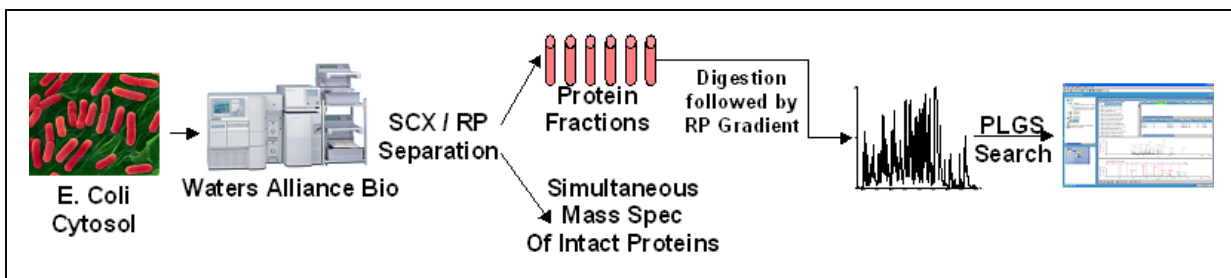*Figure 2. Schematic Representation of Off-line 2D LC/MS/MS.*

Figure 3. Schematic Representation of Protein Prefractionation Experiment.

Sample

200 μL of 50 mg/mL *E. coli* cytosol in 20 mM Ammonium Bicarbonate was diluted with 400 μL 20% acetic acid. The sample was then centrifuged at 13000g for 30 minutes and the supernatant used for injection. A total of 8 mg sample was injected onto the system.

• Salt concentration was increased in a step-wise manner and a short reversed phase separation run after each step
• Salt steps: 100 mM, 200 mM, 300 mM, 500 mM and 1 M
• RP gradient: 0 to 60% acetonitrile in 5 minutes
• The reversed phase column eluent was coupled to a 2487 UV detector and then into a WFC III fraction collector. The fraction collector was set up to collect every 90 seconds and each fraction dried down prior to digestion
• A portion of the flow was also directed into the NanoFlow™ source of a Micromass® Q-Tof micro™ mass spectrometer to enable simultaneous MS detection
• Each of the collected fractions was reduced and alkylated and then digested using trypsin in the presence of 0.1% RapiGest SF solution. The amount of trypsin used for each digestion was estimated using the mass spectrometer TIC trace.
• 10μL aliquots of each fraction were diluted three times and 1 μL injected onto the nanoscale LC/MS/MS system
• RP gradient: 5 to 30% acetonitrile in 90 minutes

Nanoscale LC System

• HPLC system: Waters CapLC and Stream Select Module
• Columns: Waters Symmetry C_{18} trapping cartridge, 0.32 mm ID x 5 mm long and Waters NanoEase Atlantis, 75 μM micron ID x 150 mm long
• Eluent flow: Approx 300 nL/min
• Column coupled to NanoLC sprayer

Mass Spectrometry

• Instrument: Micromass Q-Tof Ultima™ API equipped with NanoLockspray™
• NanoLockSpray reference solution: Glu-Fibrinopeptide B (accurate mass of doubly charged species is 785.8426 amu) and erythromycin (accurate mass of singly charged with loss of water is 716.4585 amu)
• Mass Spectrometer acquisition mode: Data Directed Analysis™ whereby species above a set intensity and with charges of 2+ and above are selected for MS/MS. A maximum of 5 co-eluting species could be selected during one decision-making process.

Data Processing

• All data was processed using ProteinLynx™ Global SERVER 2.1.
• *E. coli* species-specific database (www.ncbi.nlm.nih.gov)
• Data and results from each fraction in the on-line and off-line 2D-LC experiments were merged to give a total number of non-redundant proteins identified in each experiment
• Merge was not used in the protein pre-fractionation experiment

RESULTS AND DISCUSSION

The results obtained from the three strategies are summarized in Figure 4. Merging the data from each individual fraction in the peptide level 2D-LC separations and examination of these data led to a total number of 162 proteins identified for the on-line and 192 proteins for the off-line experiment. The data obtained from each reversed phase analysis originating from the protein level fractionation experiment was collapsed and a non-redundant list of proteins was computed. This resulted in the identification of 154 proteins. The total number of unique proteins identified from all three experiments was 318. Interestingly, even though the number of proteins identified in the protein-level experiment is the lowest at only 154, this experiment yields the greatest number of unique proteins (80). This suggests that the peptide and protein level approaches could be complementary.

An example of one of the proteins identified in both the peptide and protein-level separations is the D-Ribose binding protein, accession number 996108. In the off-line peptide separation, the protein identification results from the sequences of seven peptides and in the protein-level separation, ten peptides were correctly sequenced. Of these peptides, five were common between the experiments. The sequence of the protein with the identified peptides highlighted is shown in Figure 5. The MS/MS spectrum from one of the common peptides is shown in Figure 6.

 It is interesting to note that separating at the intact protein level and monitoring the protein separation by ESI-MS can provide additional information which cannot necessarily be obtained by peptide level, so called "bottom up" strategies.
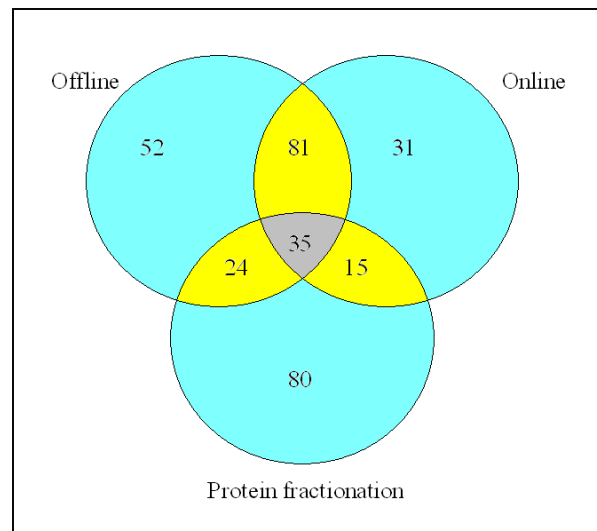

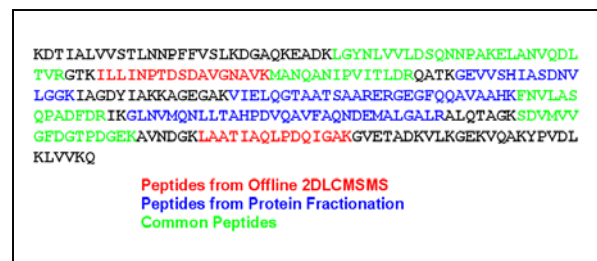Figure 4. Proteins Identified in All Strategies.

Figure 5. Sequence of Common Protein with Observed Peptides Highlighted.

An example of this is shown in Figure 7, where the ESI-MS spectrum obtained at a retention time of 19 minutes from the first salt concentration/reverse phase separation.
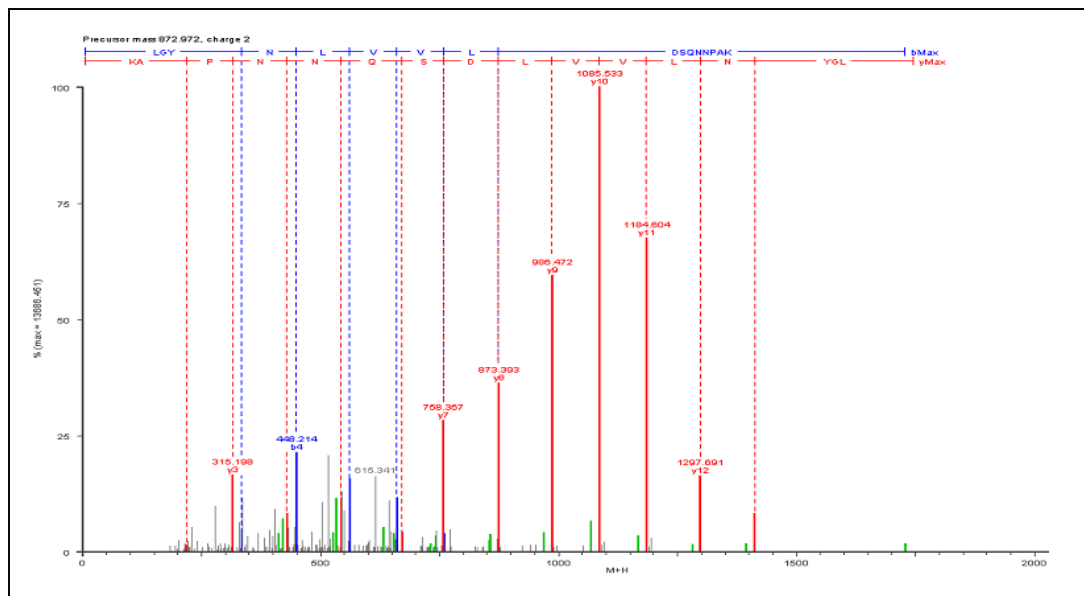
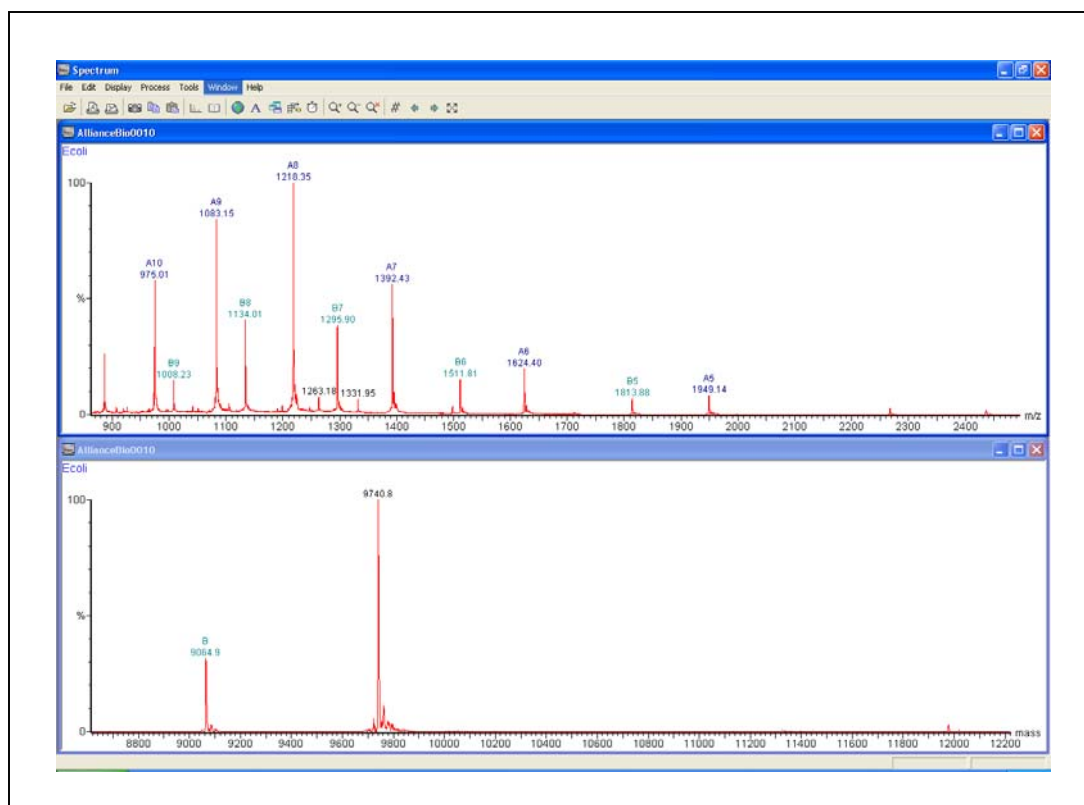*Figure 6. MS/MS Spectrum of Common Peptide, LGYNLVVLDSQNNPAK.*

Figure 7. Electrospray Mass Spectrum of the Intact Protein Separation Obtained from the First Salt Fraction at a Retention Time of 19 minutes: Multiply Charged Envelope and Maximum Entropy Processed.

The mass spectrum shows two distinct multiply charged series, and when processed by Maximum Entropy these correspond to masses at 9064.9 and 9740.8. After tryptic digestion and analysis of the same fraction by LC/MS/MS, two proteins were identified that had very similar theoretical molecular weights to those observed in the ESI-MS spectrum; Hypothetical protein yhcN precursor *E.coli* strain K12 (molecular weight 9196.4) and Chain F, Crystal Structure Of *E. coli* Periplasmic Protein Hdea (molecular weight 9740.9).

The theoretical sequence of the protein Chain F, Crystal Structure Of *E.coli* Periplasmic Protein Hdea matched to the measured mass from the intact protein mass spectrum, confirming the primary amino acid sequence. However, in the case of the Hypothetical protein yhcN precursor, there is a discrepancy between the measured mass and the theoretical mass of the protein identified from the databank, (9196.4 vs 9064.9). The experimental mass obtained differs by –131 Da, which is consistent with the mass of the N-terminal Methionine residue. If this residue is removed then the theoretical mass for this protein corresponds well with the experimentally derived mass from the electrospray mass spectrum, Figure 8. Identification of the N-terminal residue would not have been possible using a "bottom up" proteomics approach, due to the proximity of multiple lysine residues at the N-terminal. This would have resulted in several small, tryptic peptides that would not have been detected using standard MS-based approaches. It is also apparent in the coverage obtained from the LC/MS/MS data that the identified peptides are from the C-terminal portion of the protein sequence, Figure 8.
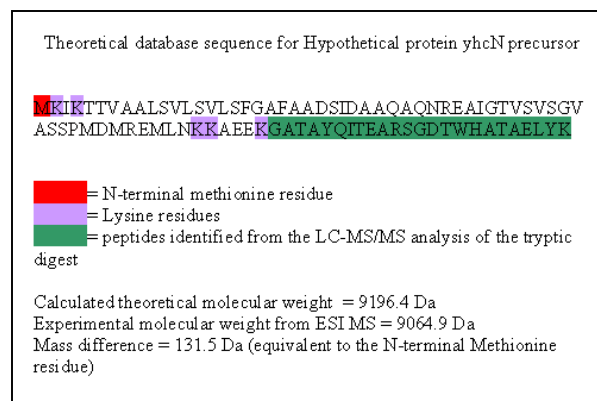

Theoretical database sequence for Hypothetical protein yhcN precursor

MKIKTTVAALSVLSVLSFGAFAADSIDAAQAQNREAIGTVSVSGV ASSPMDMREMLNKKAEEKGATAYQITEARSGDTWHATAELYK

■ = N-terminal methionine residue
■ = Lysine residues
■ = peptides identified from the LC-MS/MS analysis of the tryptic digest

Calculated theoretical molecular weight = 9196.4 Da
Experimental molecular weight from ESI MS = 9064.9 Da
Mass difference = 131.5 Da (equivalent to the N-terminal Methionine residue)

Figure 8. Sequence of the Hypothetical protein yhcN precursor E.coli strain K12. Comparison with the Experimental Mass Derived from the ESI Mass Spectrum.

SUMMARY

- Three different strategies have been compared and contrasted for the separation of an *E. coli* cytosolic sample.
- It can be summarized that the common protein identifications between all three techniques are relatively low in number, indicating that peptide and protein level analytical strategies are complementary.
- Use of the intact protein mass spectrum can yield extra information that would not be obtained using tryptic digestion in combination with mass spectrometry.

Waters

Sales Offices:

AUSTRIA 43 1 877 18 07

AUSTRALIA 61 2 9933 1777

BELGIUM AND LUXEMBOURG 32 2 726 1000

BRAZIL 55 11 5543 7788

CANADA 800 252 4752 x2205

CZECH REPUBLIC 420 2 617 11384

DENMARK 45 46 59 8080

FINLAND 358 9 506 4140

FRANCE 33 1 3048 7200

GERMANY 49 6196 400600

HONG KONG 852 29 64 1800

HUNGARY 36 1 350 5086

INDIA 91 80 2837 1900

IRELAND 353 1 448 1500

ITALY 39 02 27 4211

JAPAN 81 3 3471 7191

KOREA 82 2 820 2700

MEXICO 52 55 5524 7636

THE NETHERLANDS 31 76 508 7200

NORWAY 47 6 384 6050

PEOPLE'S REPUBLIC OF CHINA 86 10 8451 8918

POLAND 48 22 833 4400

PUERTO RICO 787 747 8445

RUSSIA/CIS 7 095 931 9193

SINGAPORE 65 6278 7997

SPAIN 34 93 600 9300

SWEDEN 46 8 555 11 500

SWITZERLAND 41 62 889 2030

TAIWAN 886 2 2543 1898

UK 44 208 238 6100

US 800 252 4752

WATERS CORPORATION
34 Maple St.
Milford, MA 01757 U.S.A.
T: 508 478 2000
F: 508 872 1990
www.waters.com

Waters

For Complete ⚓ Confidence